

# **EDA Steam**

**Adrian Bernabeu Notari**

**Análisis de una muestra del catálogo de Steam  
con juegos de entre 1997 y 2019**

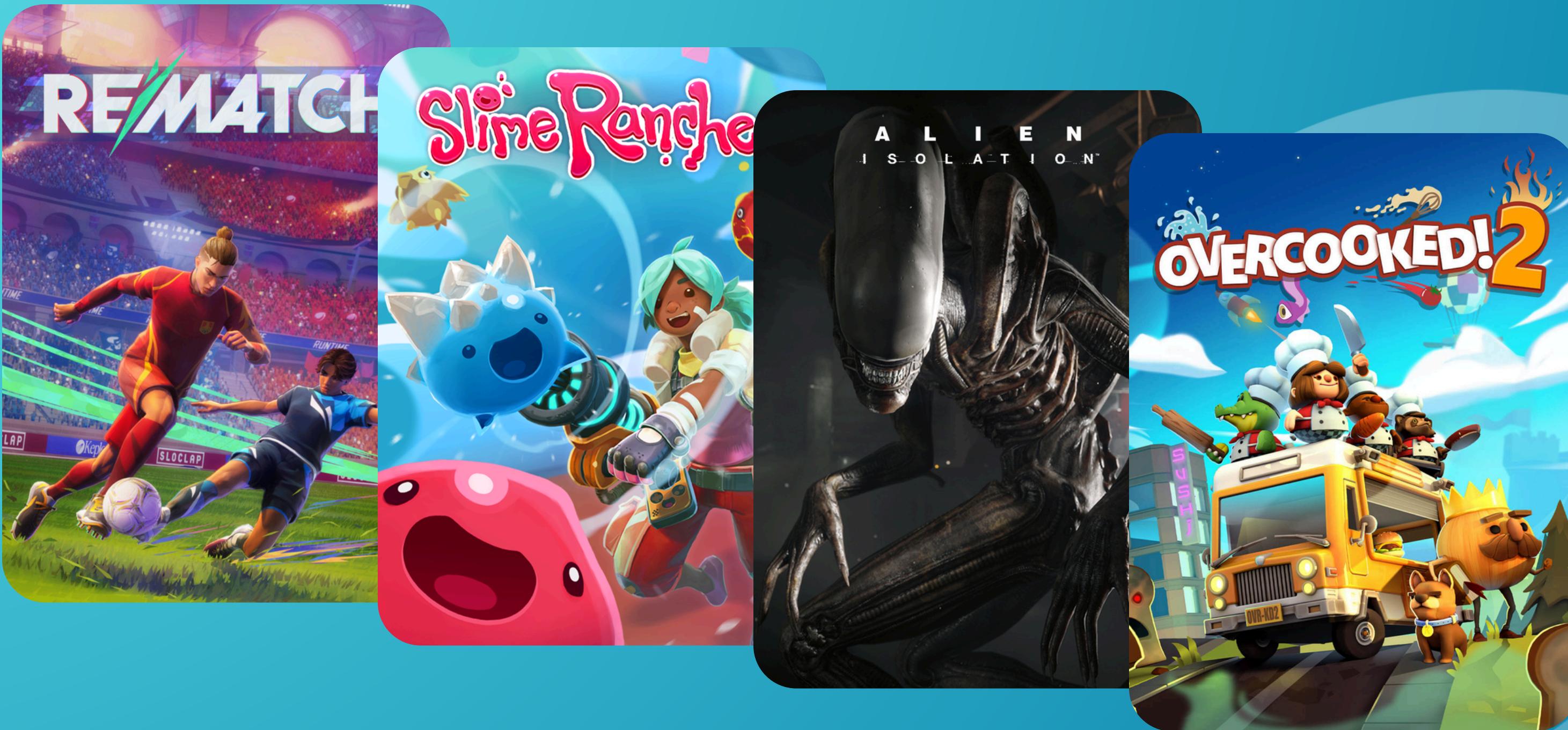


**STEAM®**



# ¿Cómo vamos a tratar los datos?

## Juegos de acción



estas categorías son confusas

# Géneros o etiquetas:

- Usamos las etiquetas en vez de los géneros
- Tenemos que 'limpiar' para quedarnos con las que nos interesen
- Mientras no se usan se obvian



Ciberpunk	+
Mundo abierto	+
Desnudos	+
Rol	+
Un jugador	+
Buena trama	+
Para adultos	+
Ciencia ficción	+
FPS	+
Futuristas	+
Primera persona	+
Ambientales	+
Exploración	+
Acción	+
Violentos	+
Gran banda sonora	+
Rol de acción	+
Aventura	+
Personalización de perso...	+
Simulador inmersivo	+



el juego se puede repetir

**Hipótesis principal:** En su mayoría los estudios suelen hacer caso de las métricas (aunque no siempre)

**Hipótesis 1:** El éxito de un juego influye en la evolución del género

**Hipótesis 2:** Un mayor precio no conlleva una mejor recepción (más caro != mejor)

**Hipótesis 3:** El precio medio de los juegos ha subido entre 1997 y 2019

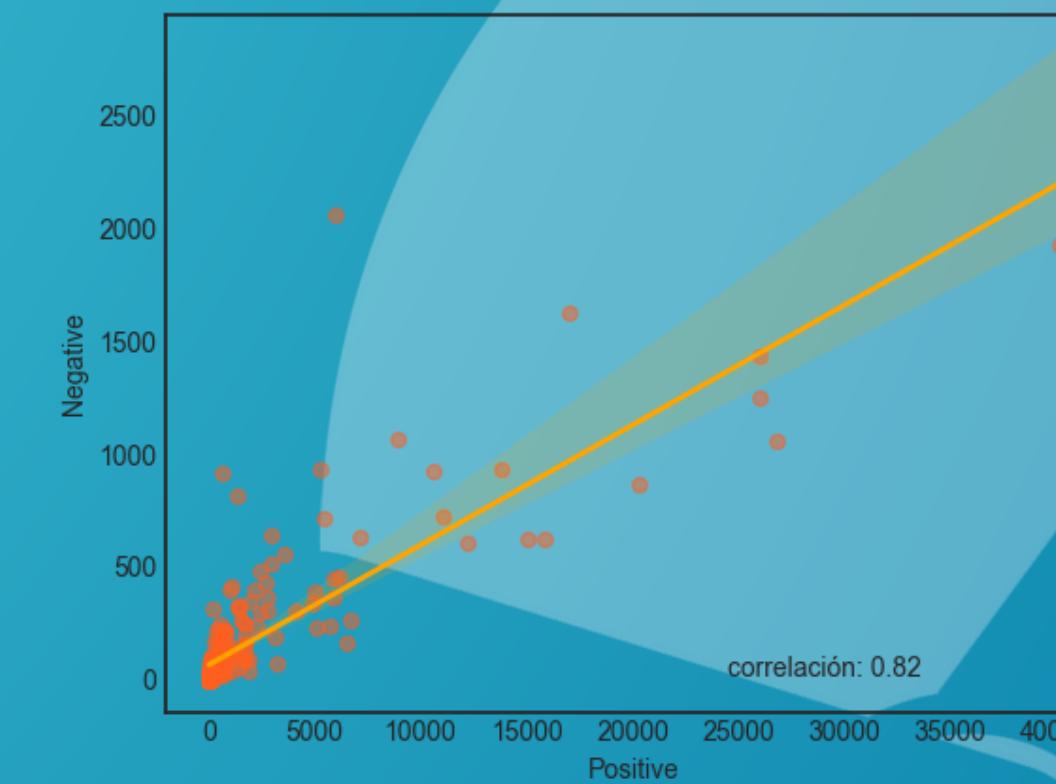
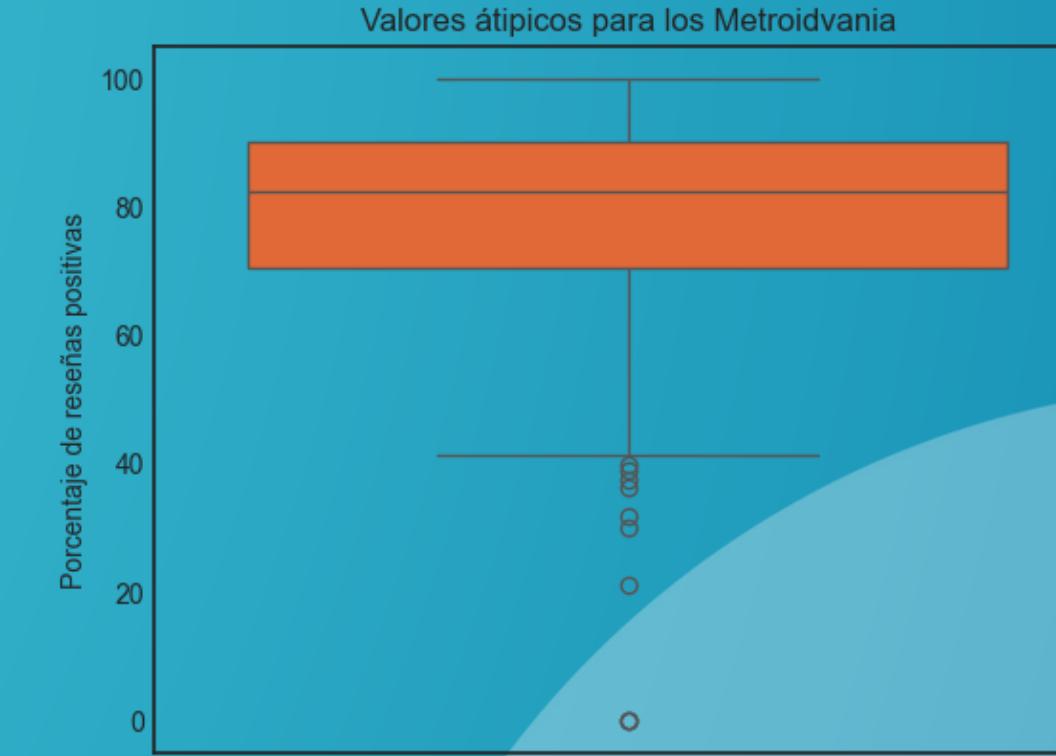
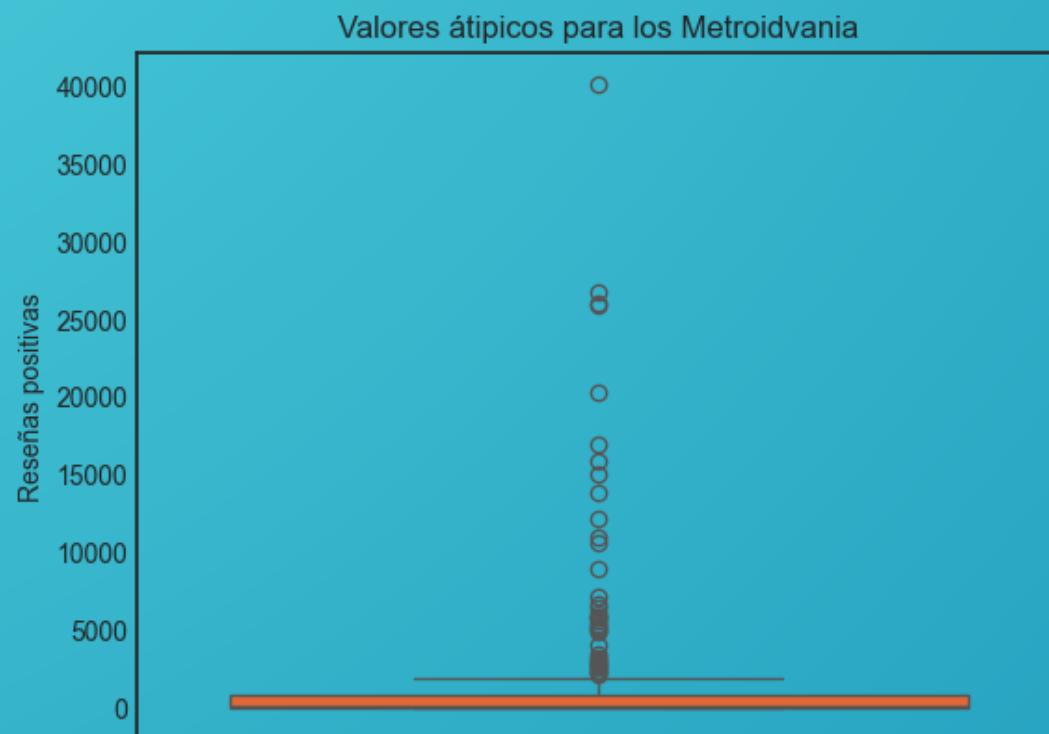
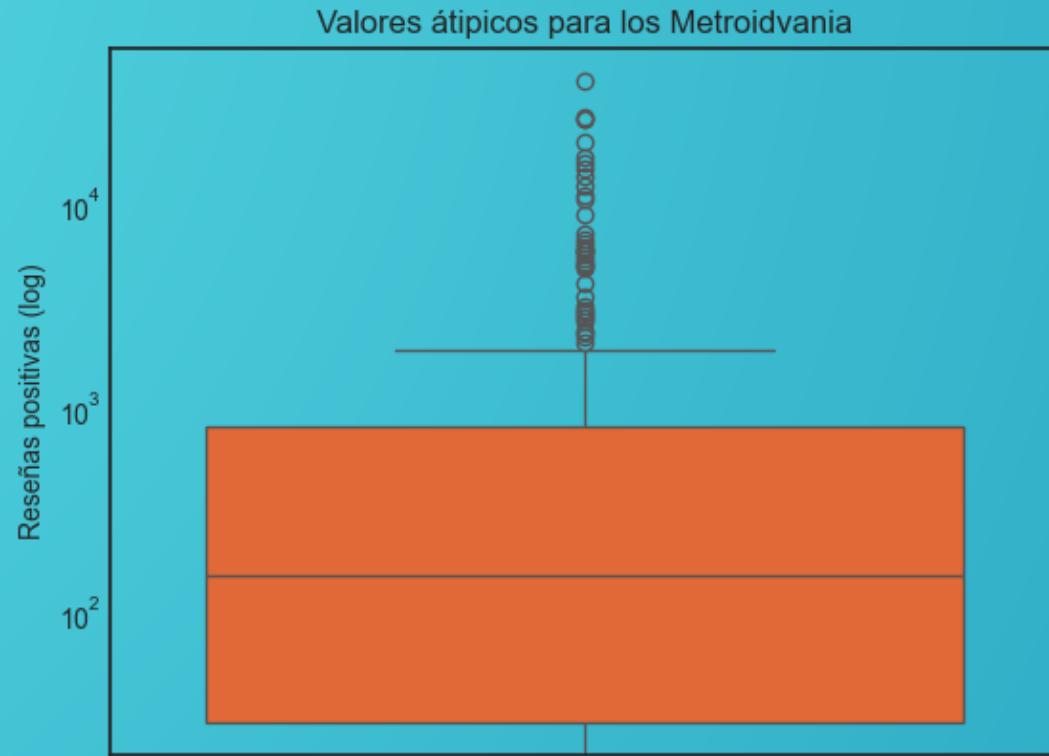
**Hipótesis 4:** La oferta no corresponde con la demanda

**Pregunta:** ¿Qué géneros se combinan más?



estructura del eda

# Hipótesis 1 Los lanzamientos exitosos influyen



lo normal es que gusten

# Hipótesis 1 Los lanzamientos exitosos influyen



Name	Positive	Release
Hollow Knight	40167	2017-02-24
Ori and the Blind Forest	26776	2015-03-10
Dead Cells	26033	2018-08-06
Batman: Arkham City - Game of the Year Edition	25995	2012-09-07
Batman: Arkham Asylum Game of the Year Edition	20297	2010-03-26

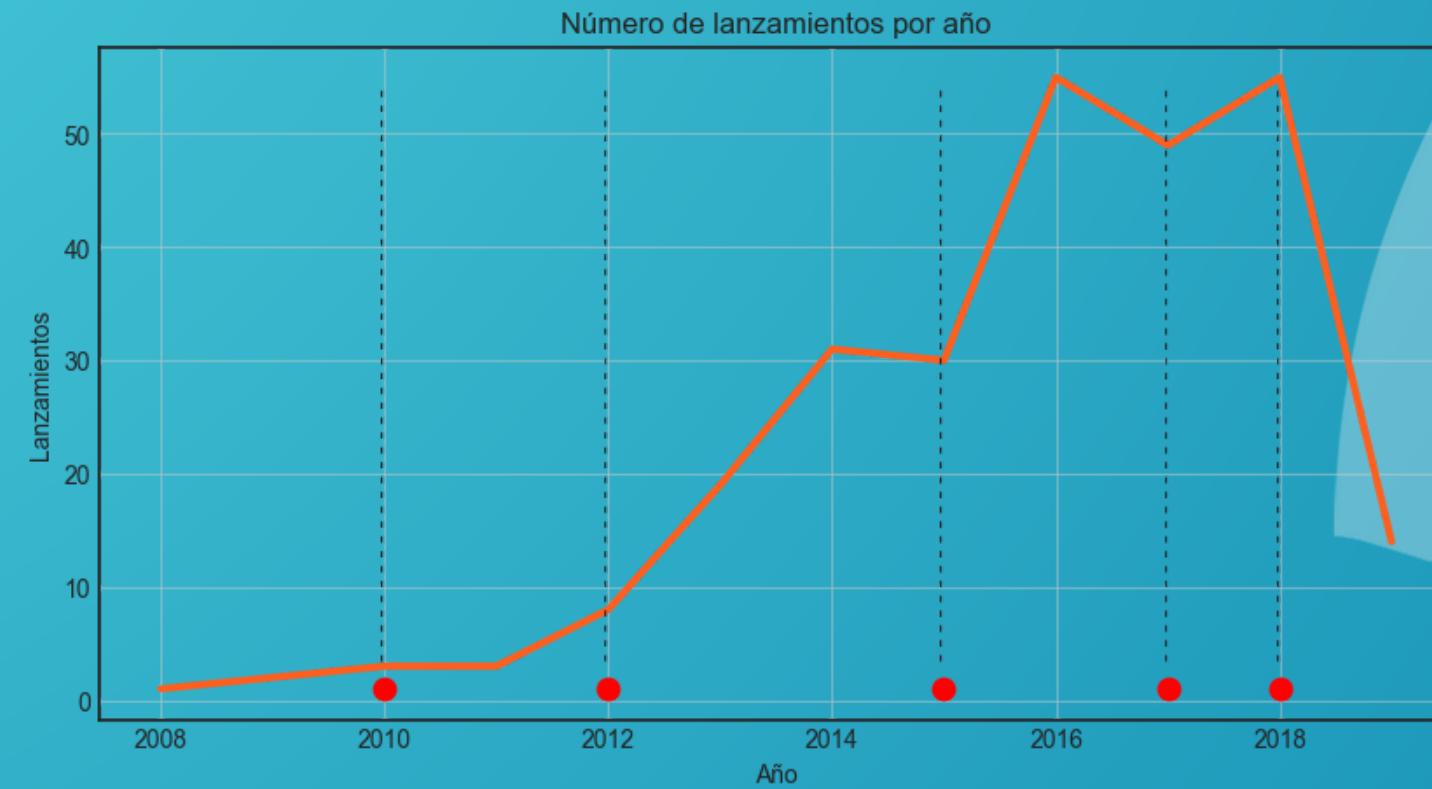


...

# Hipótesis 1 Los lanzamientos exitosos influyen

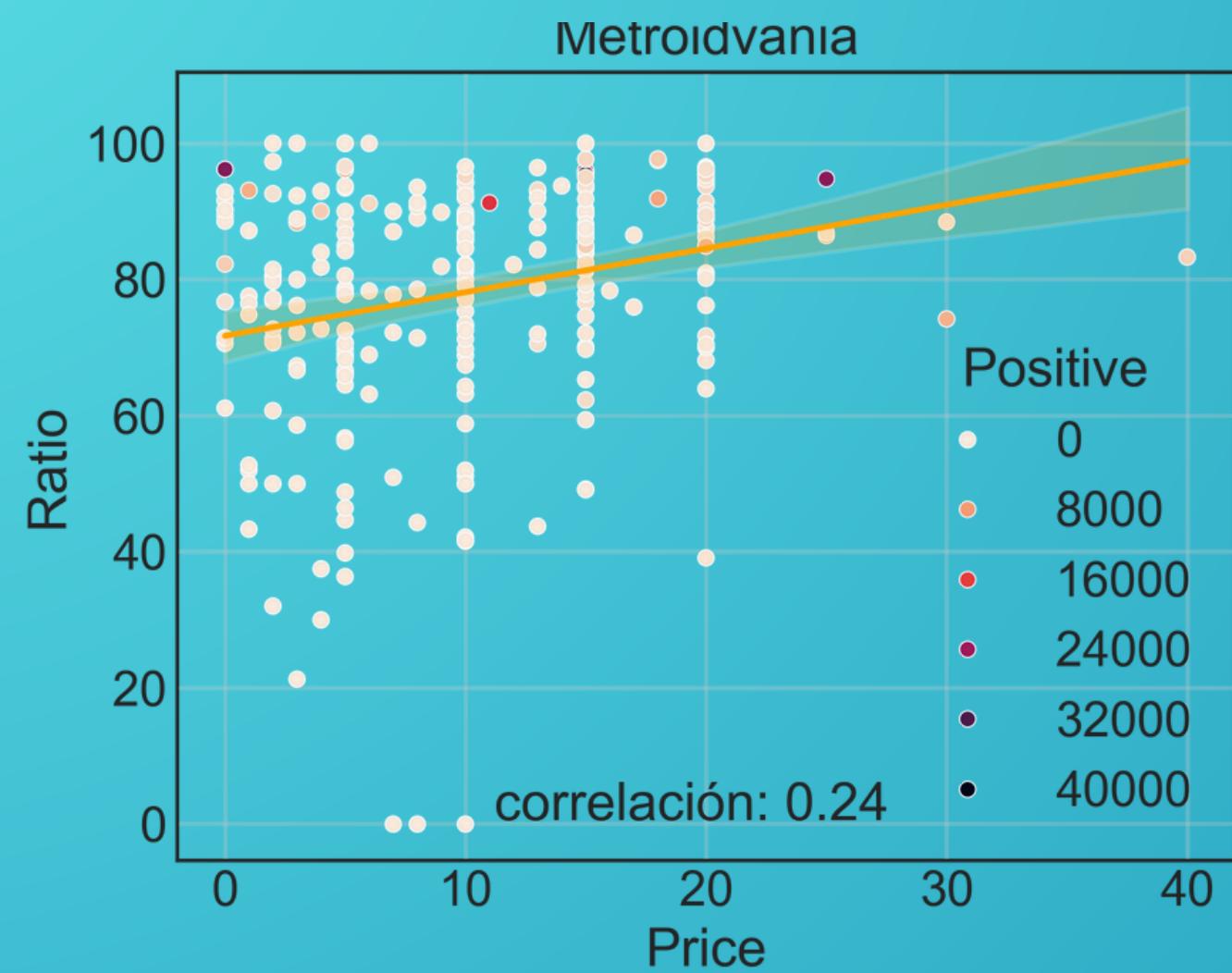


Name	Positive	Release
Hollow Knight	40167	2017-02-24
Ori and the Blind Forest	26776	2015-03-10
Dead Cells	26033	2018-08-06
Batman: Arkham City - Game of the Year Edition	25995	2012-09-07
Batman: Arkham Asylum Game of the Year Edition	20297	2010-03-26



hipótesis confirmada

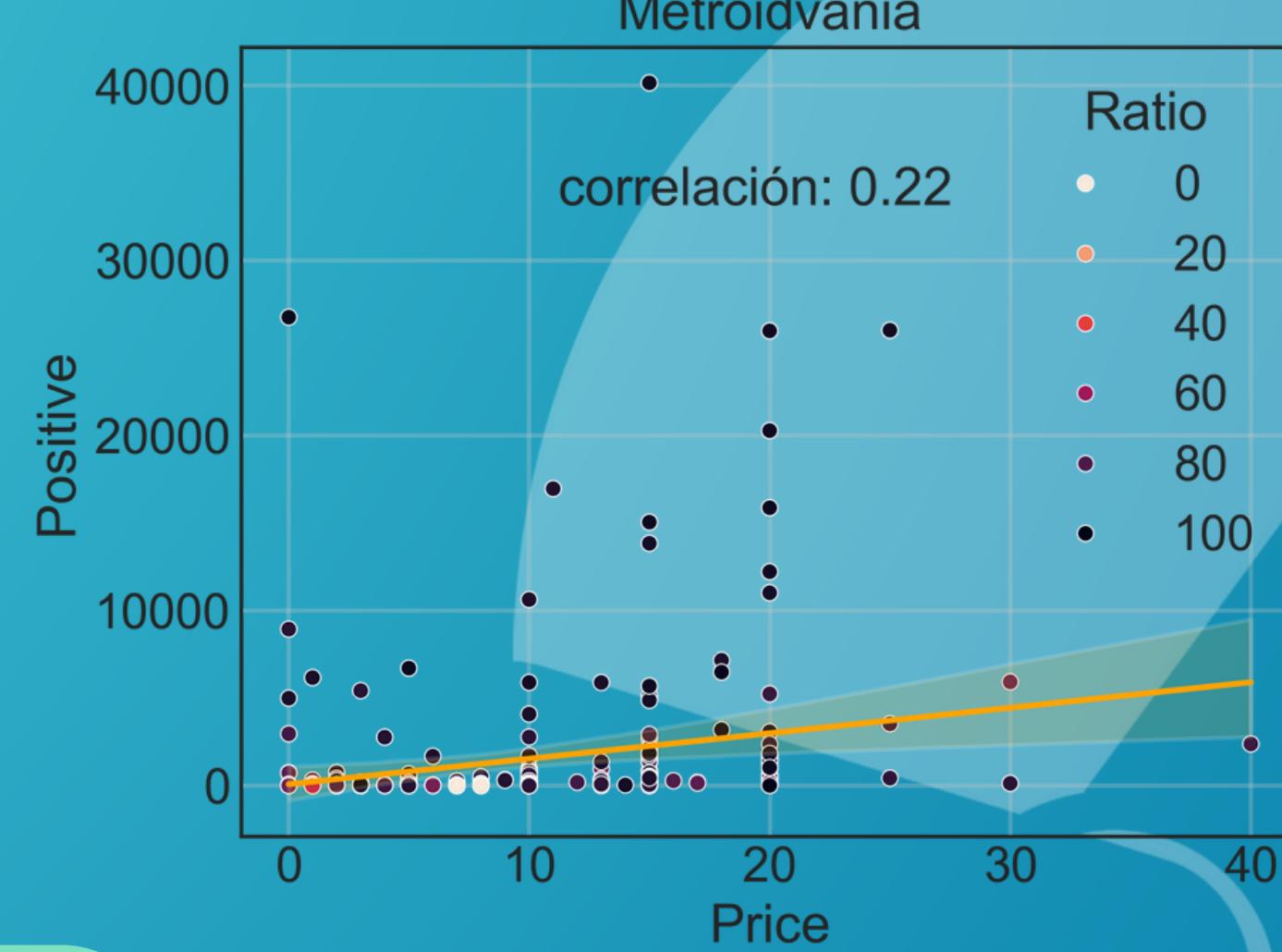
## Hipótesis 2 No por ser más caro es mejor recibido



la correlación es floja

Correlación de Pearson para:

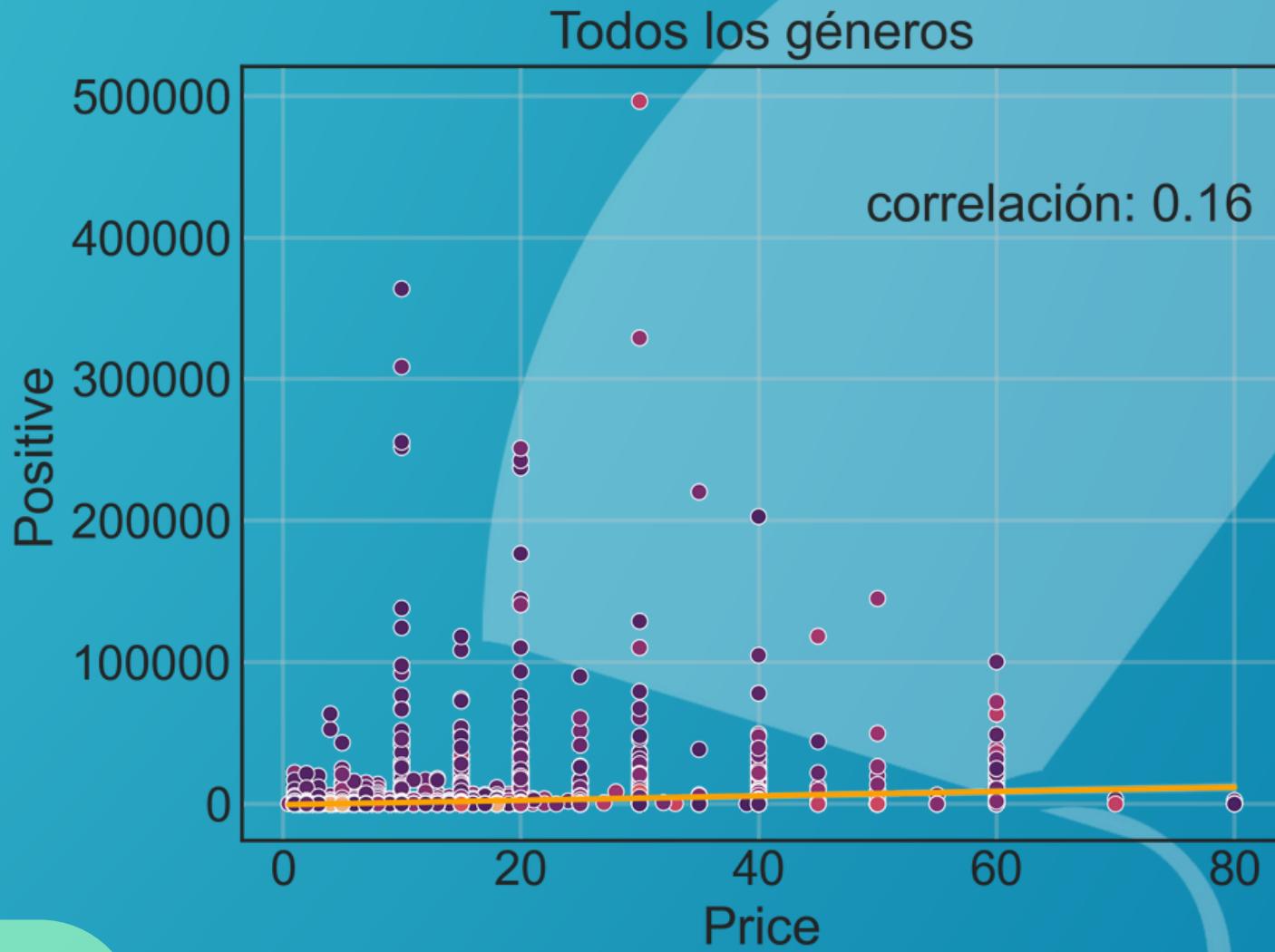
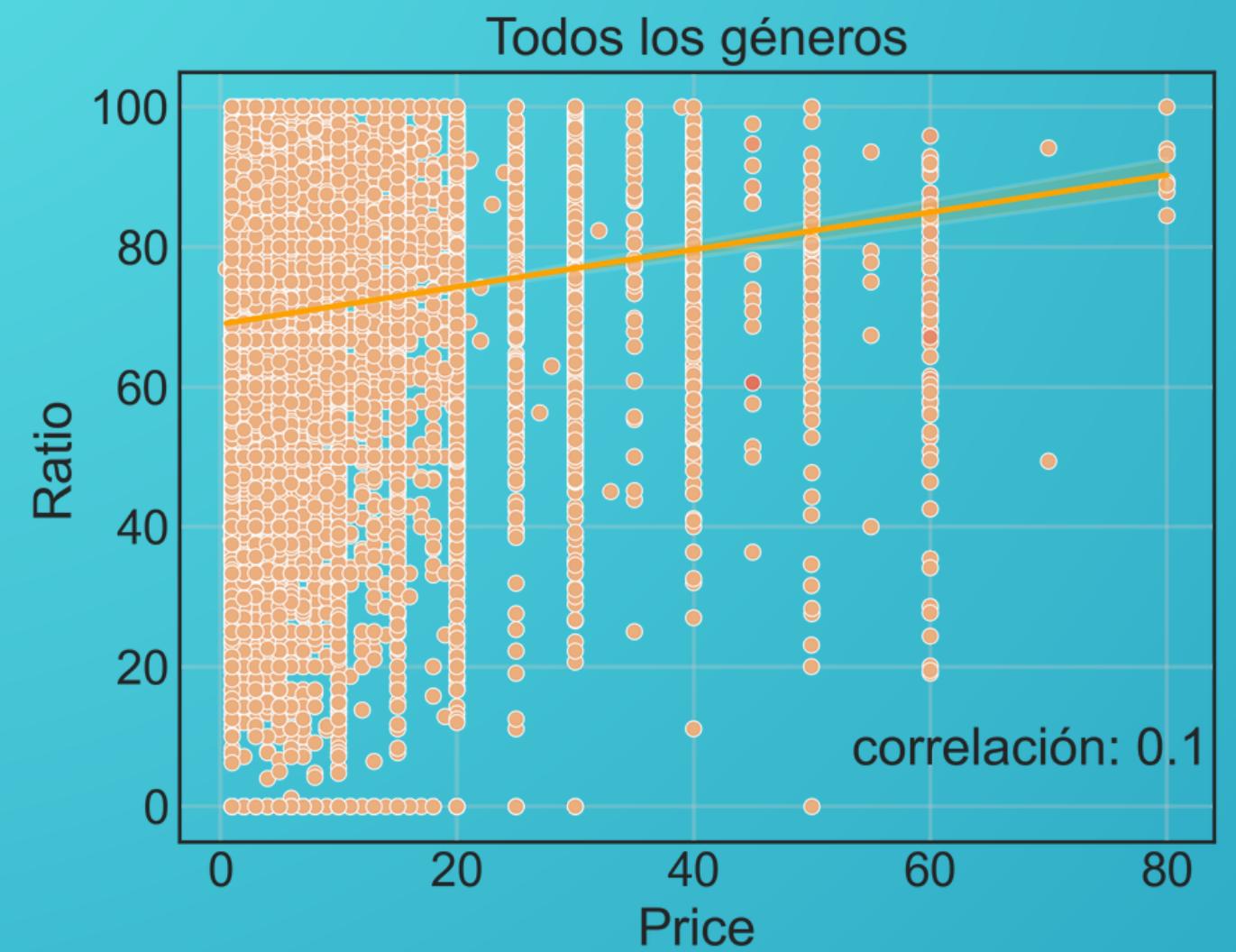
- ratio de reseñas positivas vs precio
- cantidad de reseñas positivas vs precio



## Hipótesis 2 No por ser más caro es mejor recibido

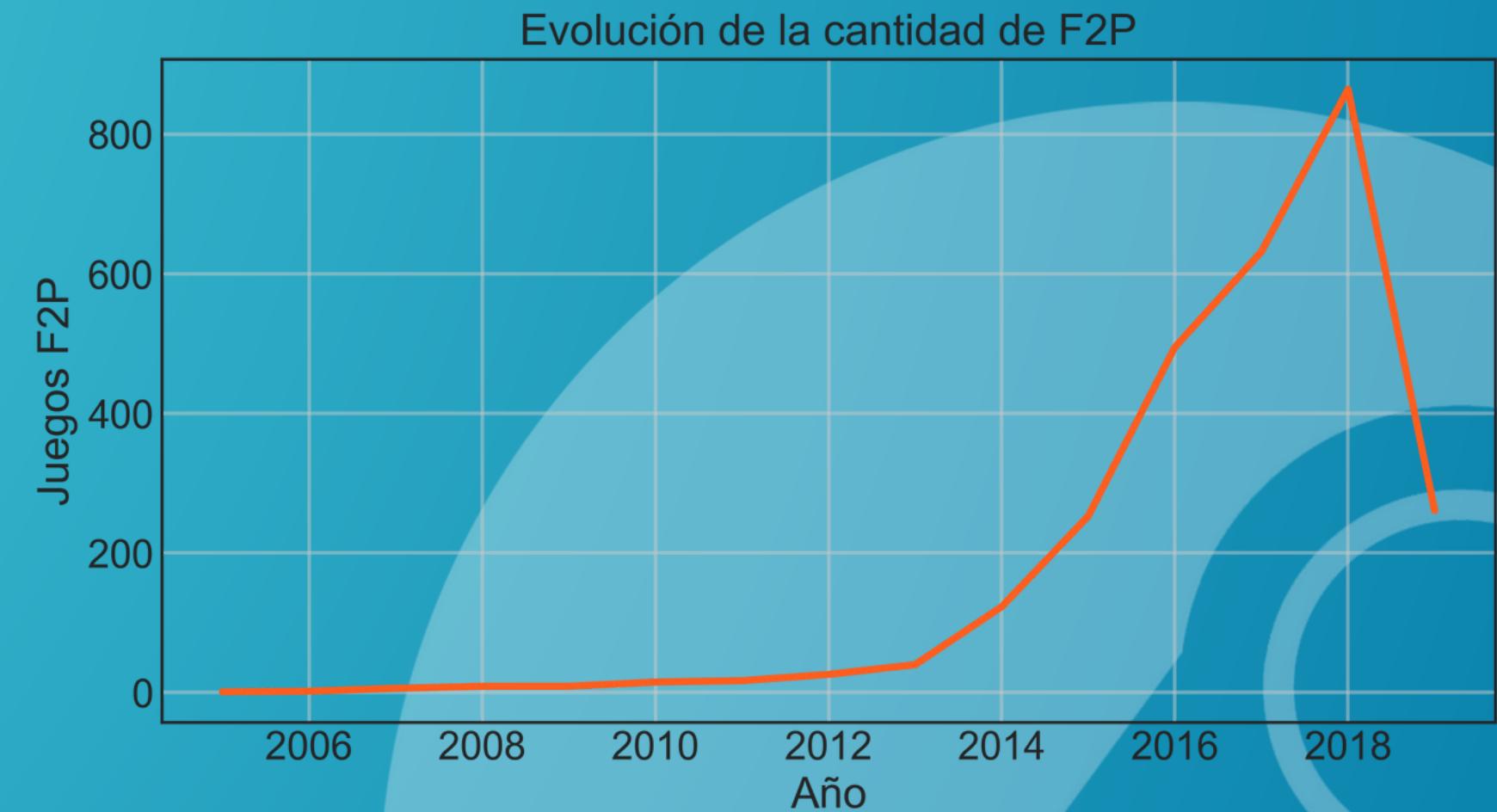
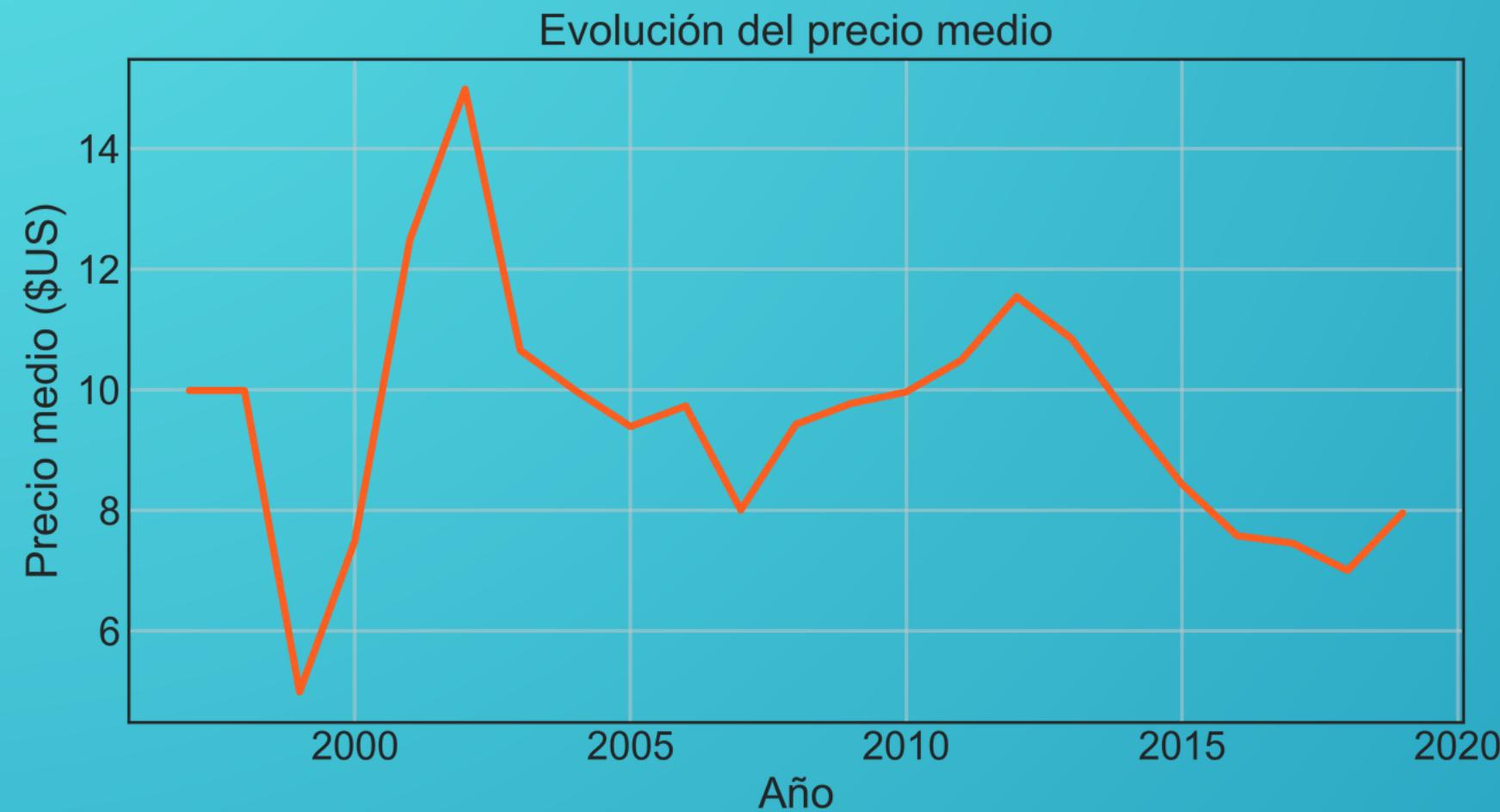
Correlación de Pearson para:

- ratio de reseñas positivas vs precio
- cantidad de reseñas positivas vs precio



hipótesis confirmada

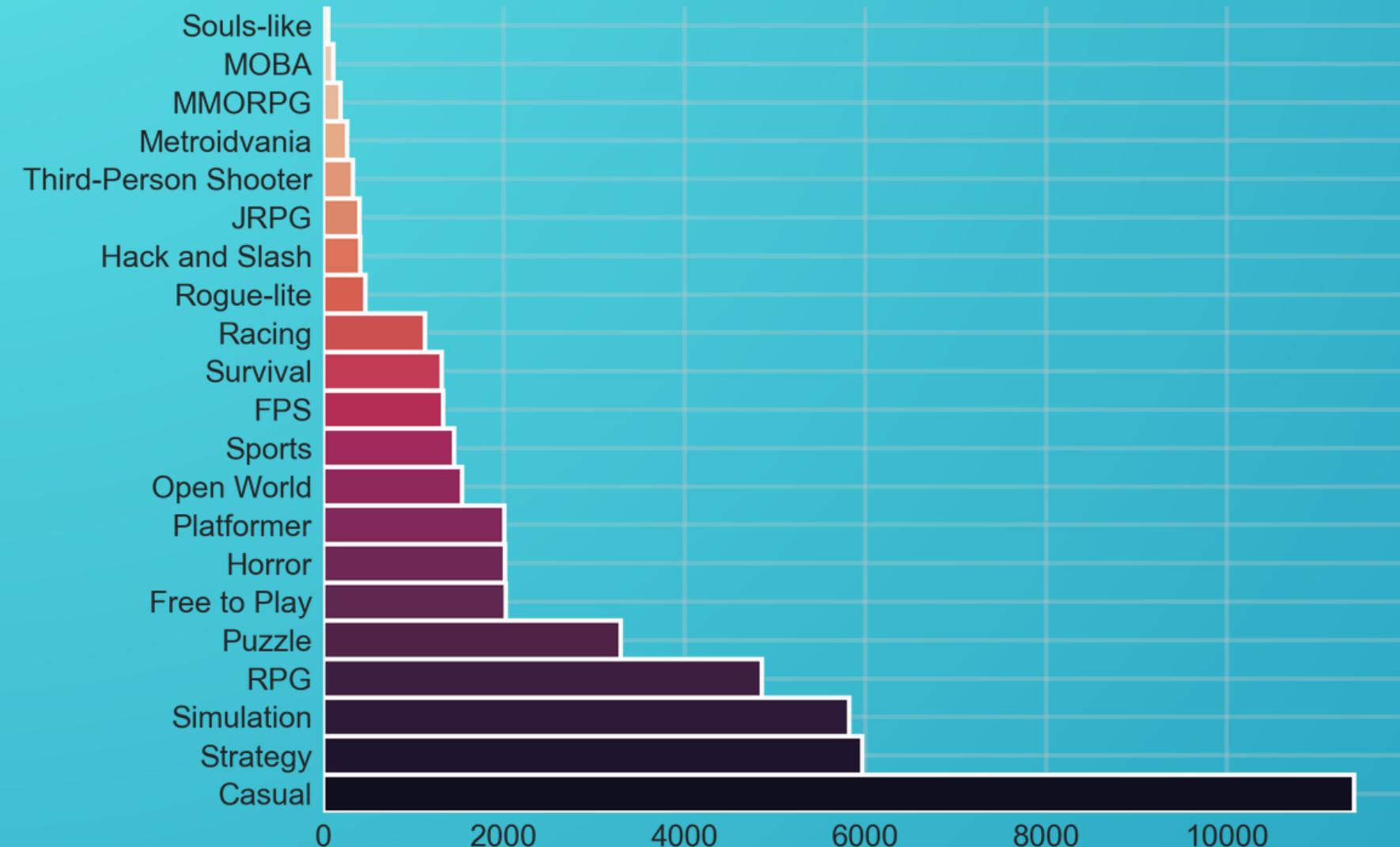
# Hipótesis 3 El precio medio de los juegos ha subido



hipótesis incorrecta

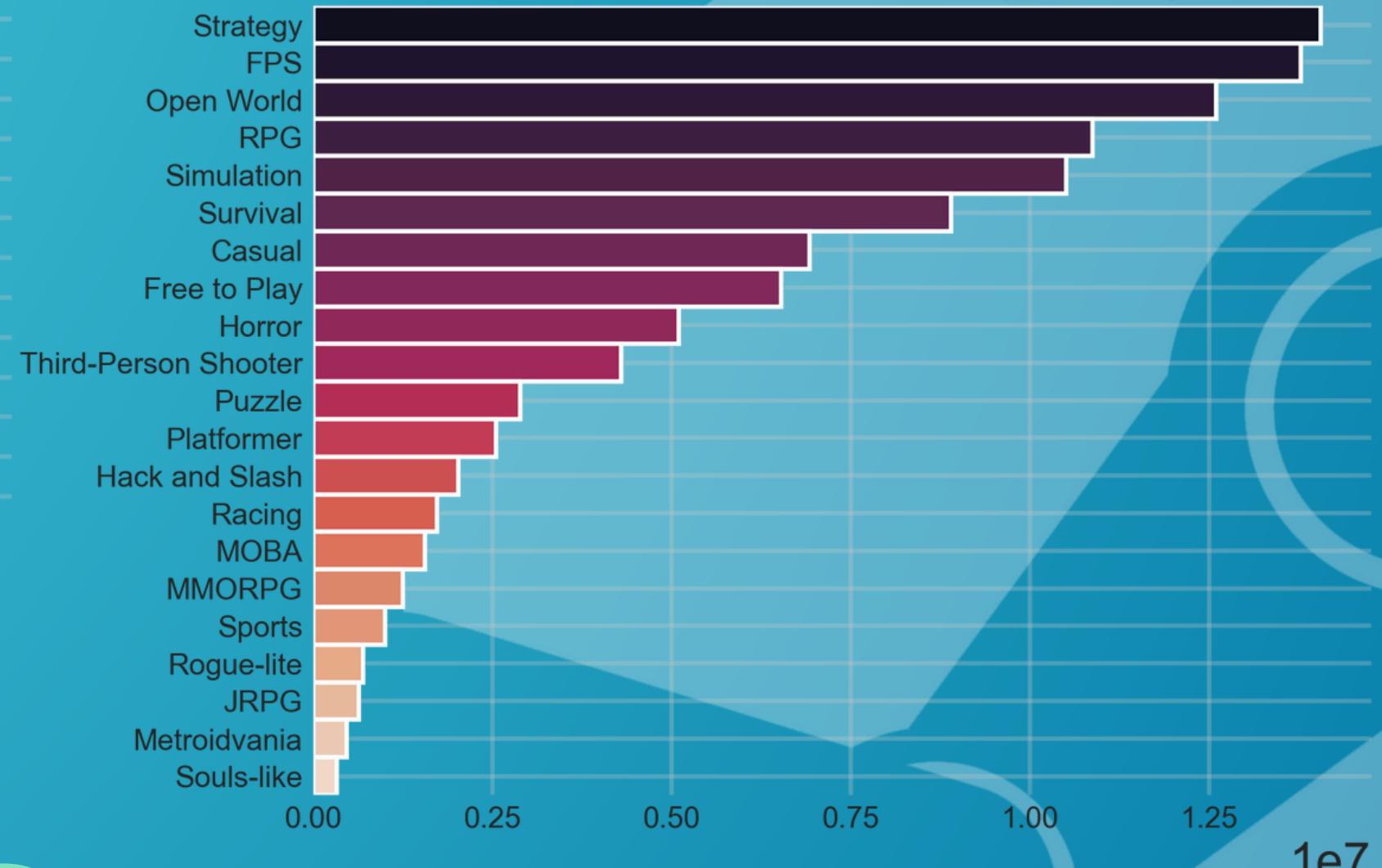
# Hipótesis 4 La oferta no corresponde con la demanda

Presencia de las etiquetas en el df



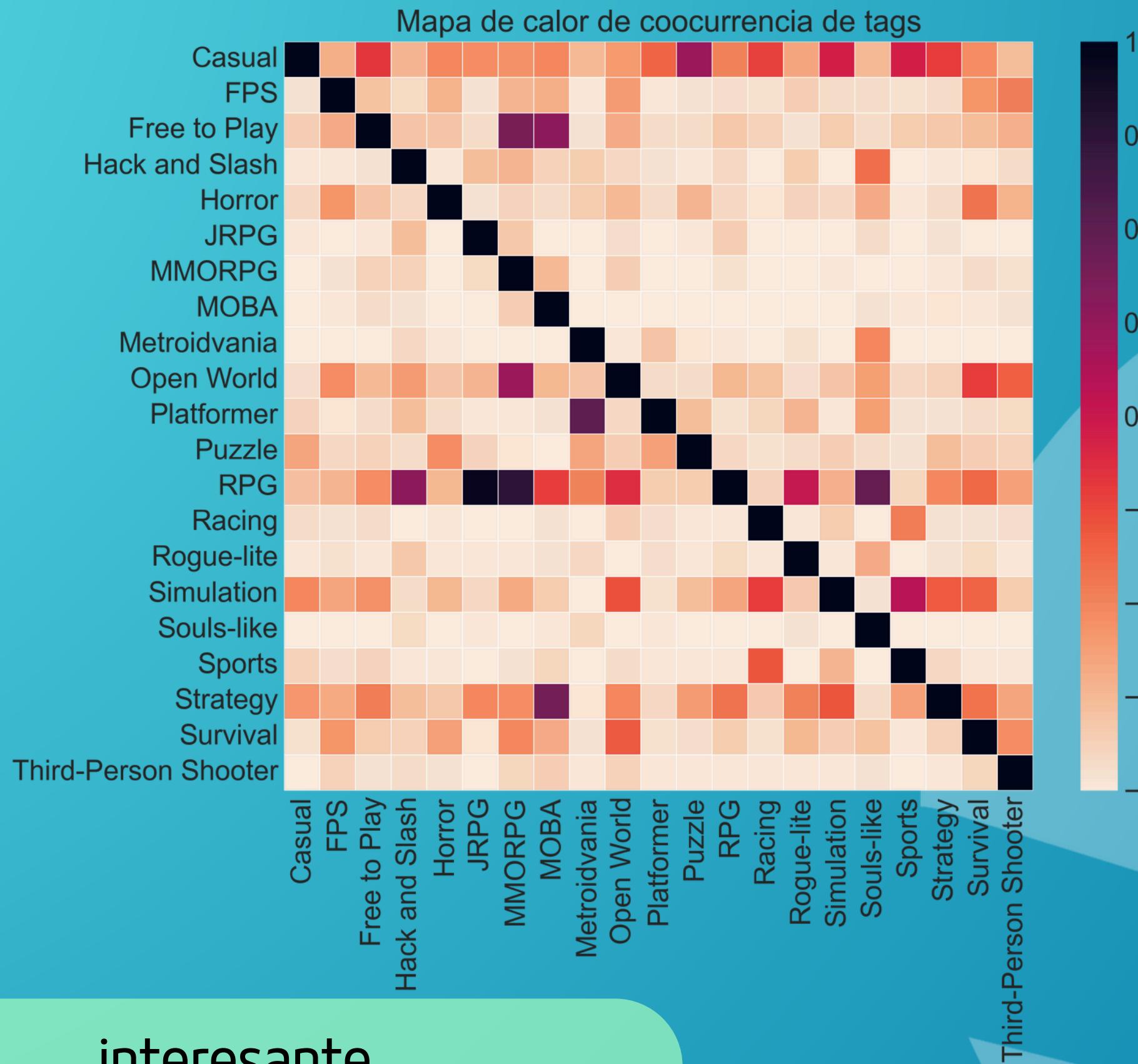
no queda claro

Número total de reseñas por etiqueta



1e7

# Pregunta ¿Qué etiquetas son las que más se combinan?



interesante

# Fin

**Hipótesis principal:** En su mayoría los estudios suelen hacer caso de las métricas (aunque no siempre)

- No determinado

**Hipótesis 1:** El éxito de un juego influye en la evolución del género

- Correcto

**Hipótesis 2:** Un mayor precio no conlleva una mejor recepción (más caro != mejor)

- Correcto

**Hipótesis 3:** El precio medio de los juegos ha subido entre 1997 y 2019

- Incorrecto

**Hipótesis 4:** La oferta no corresponde con la demanda

- No determinado

**Pregunta:** ¿Qué géneros se combinan más?

- Esto era por curiosidad



muchas gracias